

## N O T I C E

THIS DOCUMENT HAS BEEN REPRODUCED FROM  
MICROFICHE. ALTHOUGH IT IS RECOGNIZED THAT  
CERTAIN PORTIONS ARE ILLEGIBLE, IT IS BEING RELEASED  
IN THE INTEREST OF MAKING AVAILABLE AS MUCH  
INFORMATION AS POSSIBLE

ASA-CR-168748) THE CHEMICAL BASIS FOR THE  
ORIGIN OF THE GENETIC CODE AND THE PROCESS  
OF PROTEIN SYNTHESIS Semiannual Report  
(Alabama Univ. in Birmingham.) 19 p  
HC A02/MF A01

N82-21868

Unclass  
09508

CSCL 06C G3/51

THE CHEMICAL BASIS FOR THE ORIGIN OF THE GENETIC  
CODE AND THE PROCESS OF PROTEIN SYNTHESIS

Semi-annual Report to

National Aeronautics and Space  
Administration

Grant Number NGR 01-010-001

from

The Division of Molecular Biology  
University of Alabama in Birmingham  
University Station  
Birmingham, Alabama 35294

December 22, 1981



## ABSTRACT

This is a semiannual report on a project, the basic thrust of which has been to elucidate the principles upon which the process of protein synthesis and the genetic code were established. Some time ago, the work more or less delineated itself into two broad categories of study. One involved looking for selective affinities between amino acids and nucleotides, and the other was concerned with looking at the chemical reactions of activated amino acids, and for ways in which selective affinities might be responsible for selective reactions. Our past results had already shown selectivities in affinity using a variety of techniques, although accurate binding constants were not available. The present report includes extensive work on nuclear magnetic resonance (NMR) studies of both monomer-monomer and monoamino acid-polynucleotide interactions. While we are beginning to get more accurate binding constants, principally with the hydrophobic amino acids and adenylic acid derivatives, one of the major results of this research has been the development of a new method of general utility for studying any amino acid interacting with any polynucleotide. Basically, this system involves the use of methyl esters of amino acids interacting with polynucleotides. An outgrowth of the latter study comes from the fact that, at high concentrations of the hydrophobic methyl esters, insoluble complexes form with poly A. A turbidimetric study of a series of methyl esters showed the order of binding to poly A to be  $\text{trp} > \text{phe} > \text{leu} = \text{ile} > \text{val} > \text{gly} > \text{met}$ . A preliminary study of these insoluble complexes shows a 1/1 stoichiometry and also arrived at estimates of binding constants which generally agree with those determined by NMR methods.

In other work, we have completed a study of the hydrolytic properties of a set of phenylalanyl-adenylic acid derivatives over a range of pH's from 1.5 to 10. The derivatives are phe-AMP ester, phe-AMP anhydride, N-acetyl phe AMP ester and N-acetyl phe anhydride. At pH 6 the hydrolysis rates are in decreasing order  $\text{phe-anhydride} > \text{phe-ester} > \text{N-Ac-phe anhydride} > \text{N-Ac-phe ester}$ . These data,

coupled with the earlier observations that the N-blocked anhydrides form peptides with high efficiency, and that N-blocked amino acids are easier to activate with ATP, suggests that in primitive protein synthesizing systems it was the carboxyl group of the peptide that was continually reactivated. Furthermore, we have now shown that N-acetyl phe AMP anhydride is the most susceptible of this set of derivatives to nucleophilic attack by hydroxylamine. This again suggests that primitive systems could not have operated efficiently through activation of free amino acids, but would probably activate the growing peptide.

### INTRODUCTION

The origin of living systems required many things, but principal among them was the appearance of an information storage system. This amounted to the establishment of a molecular memory of past experiences, although by itself this would not have been sufficient. A mechanism for the conversion of this record (or memory) into active principals was also required. We now know that the molecular memory was DNA and that the conversion into active principals required the synthesis of proteins based on information in the DNA. The uniqueness of the storage system and the efficient beauty of its operation is a story unto itself, but our major concern has been trying to unearth the molecular properties which lie at the basis of the establishment of the translation process: a mechanism whereby the nucleic acid storage language can be converted into the protein action language. Although we certainly have not reached the point of being able to say explicitly how the process of protein synthesis, along with its genetic coding mechanism, arose, we have learned enough to be able to say that there are explicit answers, i.e., the basis does not appear to lie shrouded within the abyss of millenia of evolutionary workings devoid of discernible molecular rationale. The answer seems to lie in two interrelated phenomena - affinity and chemistry. More exactly, it is the interrelation that is important - i.e., how does the affinity affect the chemistry? While the importance of this interrelation might seem intuitively obvious, it was only after sufficient data had been collected to support the intuition

that it became the guiding principle of our research. Once this was so, it became imperative to be able to (1) determine affinities between amino acids and nucleotides, and (2) to study chemical reactions relevant to protein synthesis in light of these affinities.

Earlier studies showed that the activation of a series of hydrophobic amino acids with ATP did tend to vary as a function of the affinity of the amino acid for ATP (1). These early estimates of affinity were not very accurate, and we now needed to move in a direction yielding more precise binding constants. The use of two techniques, NMR and the formation of insoluble precipitates, seems to be yielding these data.

### Nuclear Magnetic Resonance

Monomer-monomer. The fortunate acquisition of the services and talents of Dr. M.A. Khaled, whose expertise is in the field of NMR, has enabled us to begin a more precise determination of binding constants. The majority of our experiments so far have been carried out on a JEOL PS-100 spectrometer which operates at 100 megahertz. The basic NMR experiment involves placing magnetic nuclei in a magnetic field, where the nuclei tend to align their spin axes with the direction of the field but precess around the line of the field. This is called Larmor precession. When the precessing nuclei are perturbed with electromagnetic radiation at frequencies equal to their precessing frequencies, they absorb energy and align themselves in the opposite, higher energy direction until the energy is dissipated, at which time they again assume the lower energy orientation.

While a particular nucleus has a characteristic frequency at which it will absorb energy in this manner, the immediate environment of the nucleus determines the precise value at which the absorption takes place. Consequently, if two molecules are associating with one other, the magnetic nuclei in the molecules will either be shielded or deshielded, and will therefore require different amounts of energy for the transition than when they are not associated. By observing the

differences in energy required when the two molecules are mixed in different proportions, one can estimate their affinities for each other.

Reuben and Polk (2) had used NMR to estimate the binding constants of various amino acids for 5'-AMP. In their method, they used AMP at 0.005M and induced changes in the NMR proton peak position of adenine H<sub>8</sub> and H<sub>2</sub> by adding rather high levels of tryptophan methyl ester. Since trp associates quite strongly with AMP, the induced changes were significant. They then introduced the methyl ester of a second amino acid and determined its ability to eliminate (by competition) the effect induced by trp. By proper mathematical treatment, they could then estimate the binding constant of the second amino acid for AMP. By and large, the values seem reasonable; however, the presence of two amino acids introduces the decided complication that these molecules can interact with each other in addition to interacting with AMP. Because of this problem, we wished to establish a method with fewer complications and of more general utility, especially for studying the interactions of amino acids with heterogenous oligomers or polymers of nucleotides. Our accomplishment of this objective took place in two steps. First, we set about doing experiments in much the same manner as Reuben and Polk (2), and our results were in generally good agreement with theirs. We then reversed their experiment - i.e. with the amino acid at low concentration, we varied the AMP concentration up to quite high levels. Using this technique (at pH 6.5), we did observe the movement of peak position of the amino acid protons as recorded in Fig. 1.

Using these data and the equation developed by Reuben, we calculated the binding constants of phe, leu and ile for 5'-AMP. These binding constants are shown in Table I, along with those furnished by Reuben and Polk (2). We also carried out the NMR experiment as Reuben and Polk reported and obtained essentially the same binding constants as they. One rather serious problem developed during our experiments in which the amino acid was held constant and the mononucleotide

concentration was varied. We found that at high 5'-AMP concentrations the position of the external reference signal was changing as a function of the AMP concentration as shown in Fig. 2. Also shown is the change in the position of an internal reference, tetramethyl ammonium. The movement of these reference signals is considerable, being about 20 Hz/M upfield. Since both internal and external signals moved about the same, however, we can assume that there is a general non-specific movement of all signals, including the amino acids. However, the amino acid protons all experienced additional upfield movement, and we assume this additional movement is due to specific interactions with the AMP. Before we can be sure of the binding constants, however, we must more accurately determine the non-specific movement of the amino acid protons. Regardless, it does appear that one can determine the binding constants of amino acids for nucleotides either by holding the amino acid concentration constant and varying the nucleotide or vice versa. While Reuben's method has the disadvantage of having two amino acids present at one time, changing AMP concentration has the disadvantage of movement of the reference signal due to high concentrations of nucleotide. Nevertheless, the binding constants (Table I), while different in the two methods, show the same hierarchy.

Poly A with methyl esters of amino acids. After doing a large number of experiments with monomers on the 100 MHz instrument and a few on a 400 MHz instrument, we decided to investigate homopolynucleotides. Our basic experiment was similar to that described by Razka and Mandel (3) a number of years ago. They had shown that free amino acids added to poly A caused downfield shifts of the adenine protons, principally H<sub>8</sub>. Their study was hampered by the low solubility of trp and phe. Reuben and Polk (2) had circumvented this solubility problem by using the methyl esters of the amino acids. So we combined the two experiments and used the methyl esters of the amino acids with poly A. It was a touch of good fortune. The blocking of the carboxyl group with the methyl group renders the amino acid

a cation which can then be attracted to the phosphates on the backbone of the polynucleotides. This non-specific electrostatic binding then allows a cooperative assistance to any other type of binding. The net result is a destacking of the adenine residues causing  $H_2$  and  $H_8$  proton signals to shift downfield. Fig. 3A shows the comparative movement by some of the amino acids tested, with the greatest being due to phe and the least to gly. Fig. 3B shows how the signal moved with varying leu-Ome concentrations. These data for all amino acids (Fig. 4) were then used to estimate binding constants of these amino acids for poly A using double reciprocal plots of  $\frac{1}{\delta}$  vs.  $\frac{1}{[aa]}$ , where  $\delta$  is the change in  $H_8$  peak position in hertz and  $[aa]$  is the amino acid concentration. The values are shown in Table I. Again, phe shows the highest binding constant, with leu, ile and met being nearly the same, val being less, and gly being considerably less.

The most important aspect of these observations is that we now have a method of general utility for studying the binding of any amino acid to any polynucleotide. The advantage of using the polymeric form is that now heterogenous sequences can be studied, whereas with monomers one is limited. The use of mixtures of monomers (e.g. AMP + GMP) is possible, but interpretation of the results is not nearly so easy as with heteropolymers (e.g. poly AG).

Insoluble complexes of poly A and methyl esters of amino acid. During the course of the above experiments with poly A and methyl esters, we observed that if the concentration of amino acid is carried too high, turbidity develops due to formation of transparent droplets. These can be centrifuged down, yielding a very viscous second phase. Some exploration of this phenomenon showed that different amino acid methyl esters have different abilities to cause turbidity in the poly A as shown in Fig. 5. The order is  $\text{trp} > \text{phe} \gg \text{leu} = \text{ile} \gg \text{gly} > \text{met} = 0$ . Using the midpoints of the curves to estimate 50% complexation of poly A, we estimated the various binding constants of these methyl esters for poly A. The values are, in fact, quite in line with those from other methods, but must be regarded only as estimates.



In a preliminary experiment we have analyzed the composition of the second phase from these experiments; a stoichiometry of 1/1 was found, although further experiments are necessary to make this finding more accurate.

Because the interaction of these methyl esters with poly A causes destacking of the adenine residues, one would expect a hyperchromic effect. Therefore, the binding can be studied using Ultraviolet (UV) spectroscopy. We have begun such studies and find significant increases in  $A_{257}$  due to interaction of both phe and leu methyl esters with poly A. These studies are presently being pursued.

Future studies of the interaction of methyl esters of amino acids with polynucleotides will be on a Nicolet 300 MHz instrument which has recently been acquired by our Chemistry Department and will be available for our part-time use for a nominal charge. The beauty of being able to continue these experiments on an instrument of such high resolving power is that the individual protons of an amino acid can be observed. Consequently, we will not only be able to determine binding constants, but will also be able to determine which protons on the amino acid are in closest proximity to the nucleotide base rings, i.e. we will be able to say something about the mode of interaction.

#### Reactivity studies

1. Hydrolysis. In our continuing effort to understand the chemistry of activated aminoacyl derivatives, we have now completed a study of the hydrolysis of a set of derivatives of phenylalanine and 5'-AMP consisting of free and N-acetyl esters and free and N-acetyl anhydrides.

These compounds were all synthesized in our labs and the hydrolysis studies were carried out at 25°C in a jacketed reactor. The disappearance of active anhydride species was followed by the hydroxamate assay of Lipmann and Tuttle (4), whereas the hydrolysis of the esters was followed by the appearance of periodate-positive adjacent ribose hydroxyls. A pH range from 1.5 to 10 was covered. Data from the hydrolyses were plotted according to pseudo first order kinetics,

yielding straight lines from which the rate constants were estimated. Fig. 6 shows a plot of the rate constants for the four compounds. At pH 7, the free amino acid anhydride is far and away the most reactive, followed by the free amino acid ester, the N-acetyl anhydride, and the N-acetyl ester. The N-acetyl anhydride is the only form which shows increased hydrolysis at low pH, probably due to the elimination of the positive charge on the  $\alpha$ -amino group which then allows easier approach of  $[H_3O]^+$ .

These results are interesting in several respects. First, the free amino acid anhydride, which is the first activated form in contemporary protein synthesis, is extremely unstable. It is unlikely that it would be the active species in primitive systems because it would be lost before it could be used in peptide formation. On the other hand, the N-acetyl anhydride (a model of peptidyl anhydride), while it is quite stable to hydrolysis as shown in the present studies, is extremely efficient ( $\sim 100\%$  under some conditions, see Report, Dec. 1980) at yielding peptides using either a free amino acid or an aminoacyl amide as acceptor. Furthermore, the N-Ac-phe AMP ester (which models the peptidyl-tRNA ester in contemporary systems) must be the reactive species in forming a peptide bond, yet it seems very unreactive in the present studies.

The most reasonable scenario for a primitive system thus seems to be a continued reactivation of the carboxyl group of the growing peptide, i.e. an amino acid would only be activated after its incorporation into a peptide. This type of model is further supported by our earlier studies which showed that N-blocked (peptidyl) amino acids are easier to activate with ATP than are free amino acids. The fact that protein synthesis in primitive creatures is still initiated by an N-blocked amino acid (N-formylmethionine) is also consistent with such a model.

2. Hydroxylaminolysis. The above data concerns hydrolysis. Certainly we must be concerned with hydrolysis, since it is the major competing reaction when considering peptide bond formation. The reaction of more direct relevance to

peptide bond formation, however, has to do with susceptibility to attack by amino groups, as in protein synthesis.

As an initial approach we have begun an investigation of the reaction of the various activated forms as used in the hydrolysis study. We have studied, in particular, the reaction of these activated forms with hydroxylamine. This reagent was chosen because it has an amino group with a  $pK_a$  of 6.5, which means that it is a fairly good nucleophile, even at pH 5.5. At this pH, of course, hydrolysis is minimized. Furthermore, the reaction product itself can be quantitated by complexation with  $FeCl_3$  and reading at 495 nm.

For these experiments, the samples were incubated with 0.25 M hydroxylamine at 25°C in a jacketed reaction vessel. The pH (5.5) was constantly monitored. Samples were removed at periodic intervals and assayed for the amount of hydroxamate positive material already formed during the reaction. The amount of active species still remaining plus the amount already formed was assayed by putting a sample in 2 M hydroxylamine at pH 6.5. The latter assay showed that at pH 5.5, essentially none of these compounds was hydrolyzed. Furthermore, much to our surprise, the N-acetyl aminoacyl anhydride was more reactive than any of the other species. Figure 7 shows that the order of reactivity is N-Ac-phe anhydride > phe-anhydride = phe ester > N-Ac-phe ester.

If we now compare these data with the hydrolysis data in Figure 6, we can see that while the N-blocked anhydride is the most reactive toward  $NH_2$  groups, it is quite stable to hydrolysis. As suggested earlier, all of the data considered now point to the idea that in primitive protein synthesis, the peptide (i.e., N-blocked) carboxyl group was the one continually reactivated, rather than each amino acid being activated prior to incorporation.

#### FUTURE WORK

Our future work will continue along the lines pointed out earlier on complexation and on the chemistry of peptide bond formation as influenced by affinity. In

addition, we will study the effect of polynucleotides on the rates and efficiencies of peptide bond formation.

The NMR studies will continue using the newly developed technique of using polynucleotides with aminoacyl-O-methyl esters, especially looking at heterogenous polymers and modes of binding.

#### REFERENCES

1. Lacey, J.C., Jr. and Mullins, D.W., Jr. in Origin of Life, Y. Wolman, ed., D. Reidel Pub. Co. (1981) pp. 447-456.
2. Reuben, J. and Polk, F., J. Mol. Evol. 15: 103-112 (1980).
3. Razka, M. and Mandel, M., Proc. Natl. Acad. Sci. U.S.A. 56: 1325-1329 (1971).
4. Lipmann, F. and Tuttle, L.C., J. Biol. Chem. 159: 21 (1945).

Table I. Binding Constants of Several Amino Acids  
for 5'-AMP and Poly A

<u>Amino Acid</u>	<u>Reuben and Polk<sup>a</sup></u>	<u>UAB<sup>b</sup></u>	<u>UAB<sup>c</sup></u>	<u>UAB<sup>d</sup></u>
Phe	5.1	6.8	6.9	5.0
Leu	3.1	5.1	3.4	1.6
Ile	2.9	1.7	2.3	1.6
Met	3.0	-	2.3	-
Val	2.0	-	2.3	0.8
Gly	1.7	-	0.37	0.5

<sup>a</sup>Reuben and Polk (2)

<sup>b</sup>Low concentration of amino acid with increasing concentrations of 5'-AMP,  
as discussed in the text

<sup>c</sup>Constant poly A concentration and varying concentrations of the amino acid  
methyl esters

<sup>d</sup>Estimated from turbidity curves as described in the text for mixtures of  
poly A and the methyl esters of the amino acids

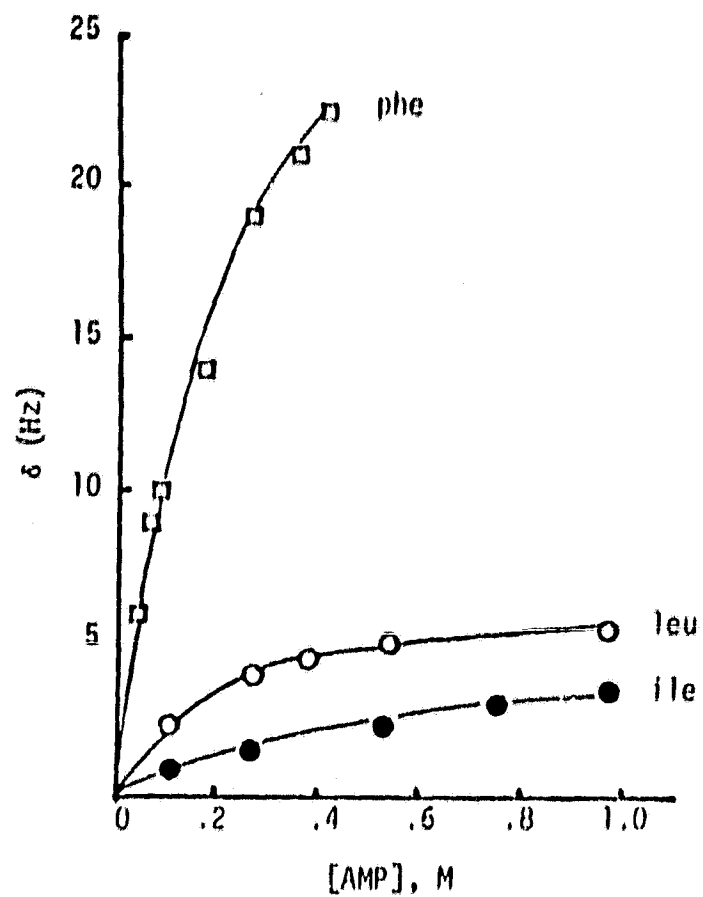


Fig. 1. Plot of the upfield change in chemical shift of phe, leu and ile protons as a function of AMP conc. See text for details.

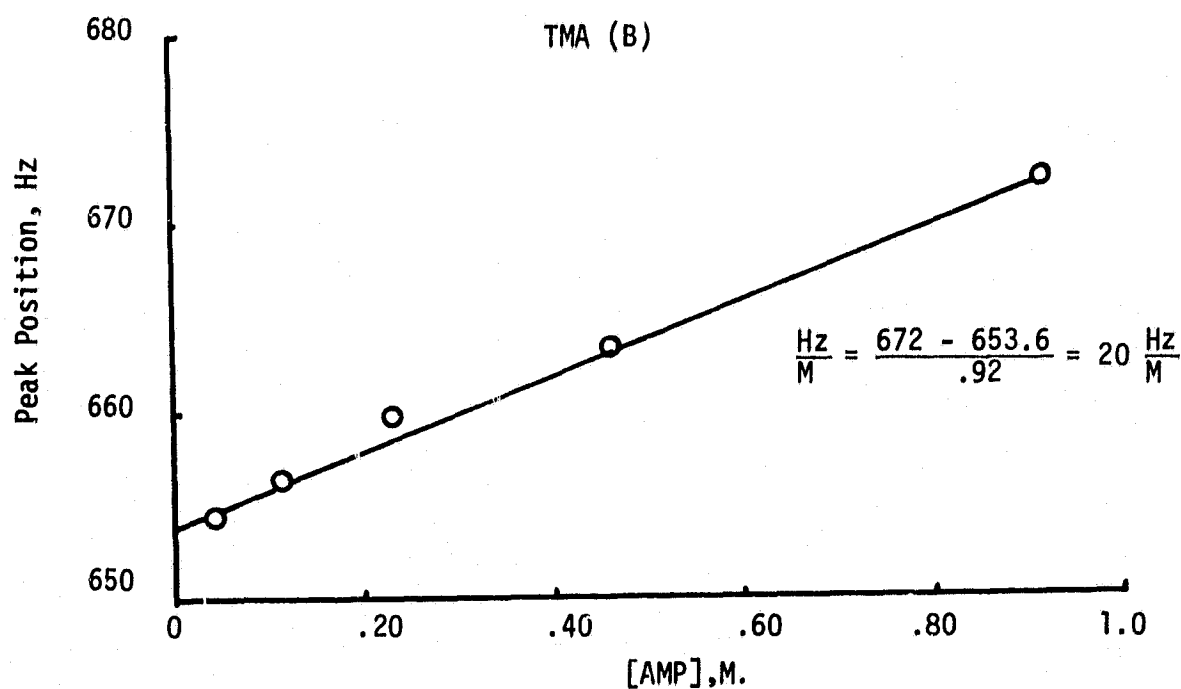
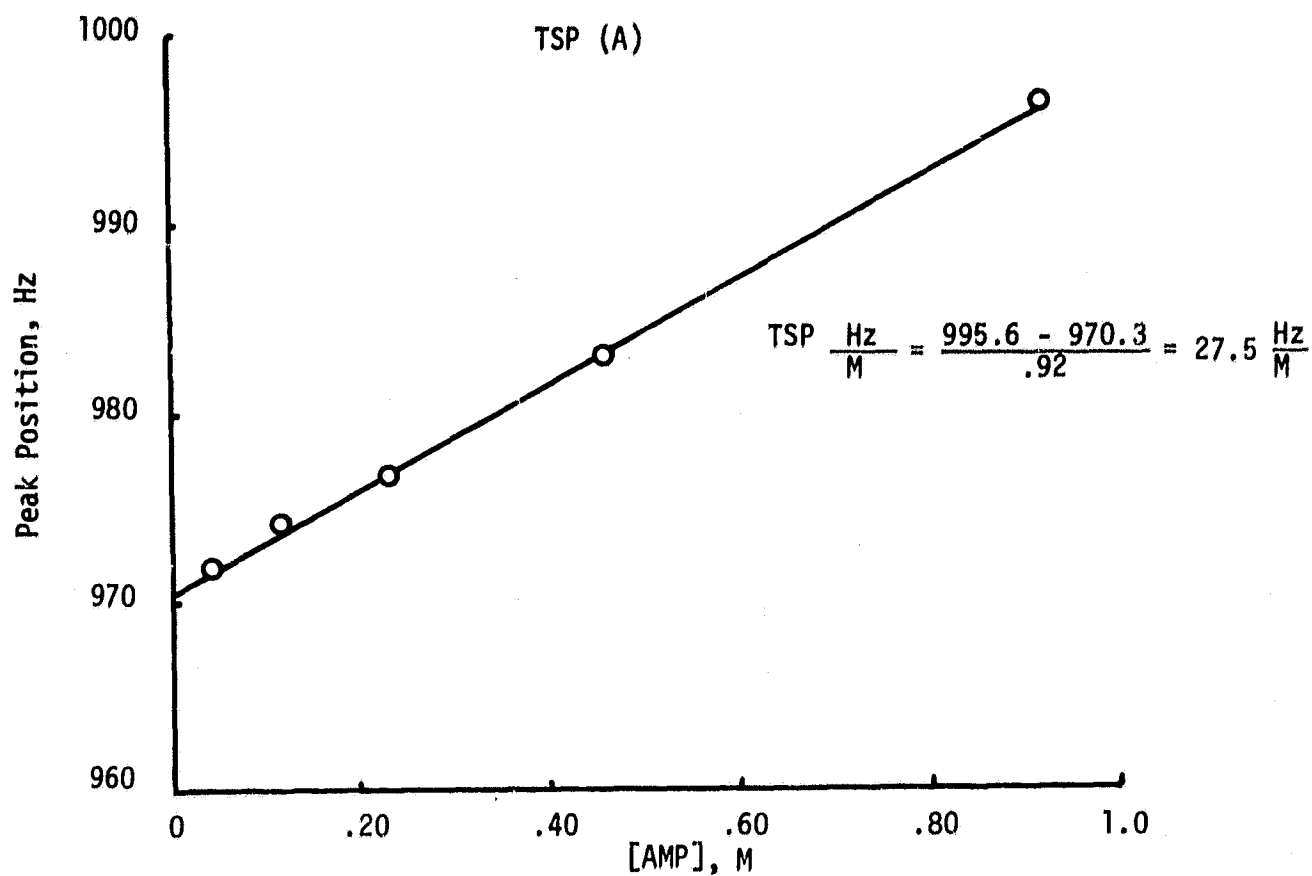


Fig. 2. Upfield movement of external TSP (A) and internal tetramethyl ammonium (B) signals as the concentration of AMP is increased. 100 MHz JEOL spectrometer.

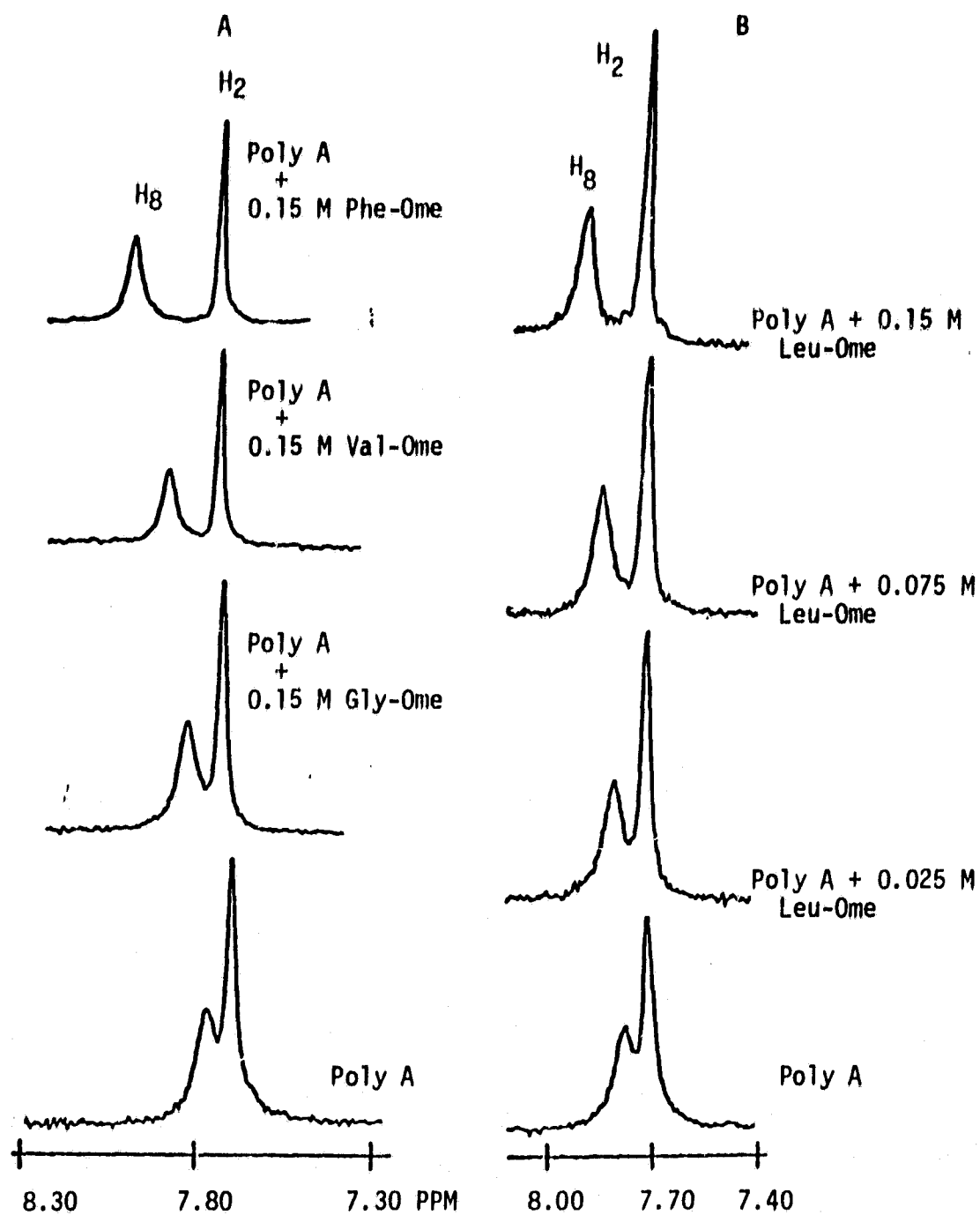
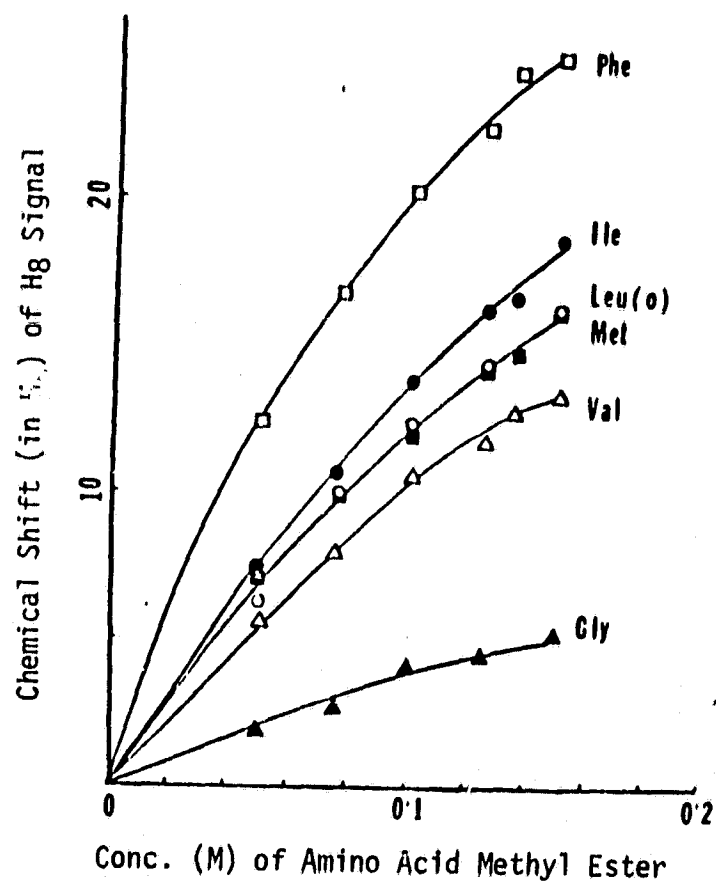


Figure 3. 100 MHz  $^1\text{H}$  NMR spectra of poly A (0.05 M) in D<sub>2</sub>O with 0.05 M phosphate buffer adjusted to pH 7.0. Shown here are the H<sub>2</sub> and H<sub>8</sub> proton resonances of the adenine ring positioned with respect to external TSP used in a capillary tube.



Figure 4. Plot of downfield shift of H $\delta$  proton signal of poly A as the function of increased concentration of amino acid methyl esters.



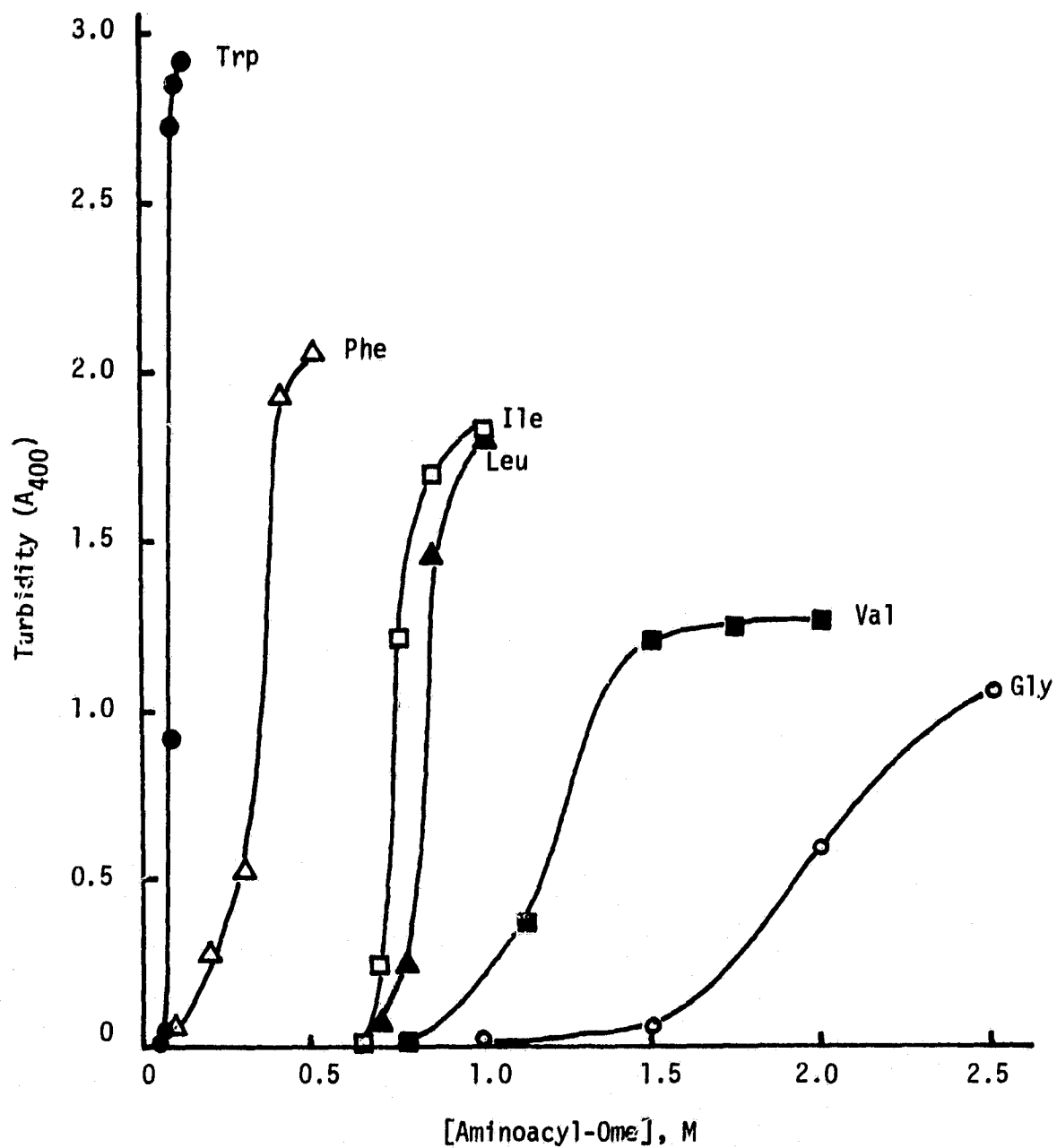


Fig. 5. Generation of turbidity (apparent absorption at 400 nm) of  $10^{-3}$  M poly A solutions in the presence of various concentrations of the methyl esters of several amino acids at pH 7.0 in 0.05 M phosphate buffer.

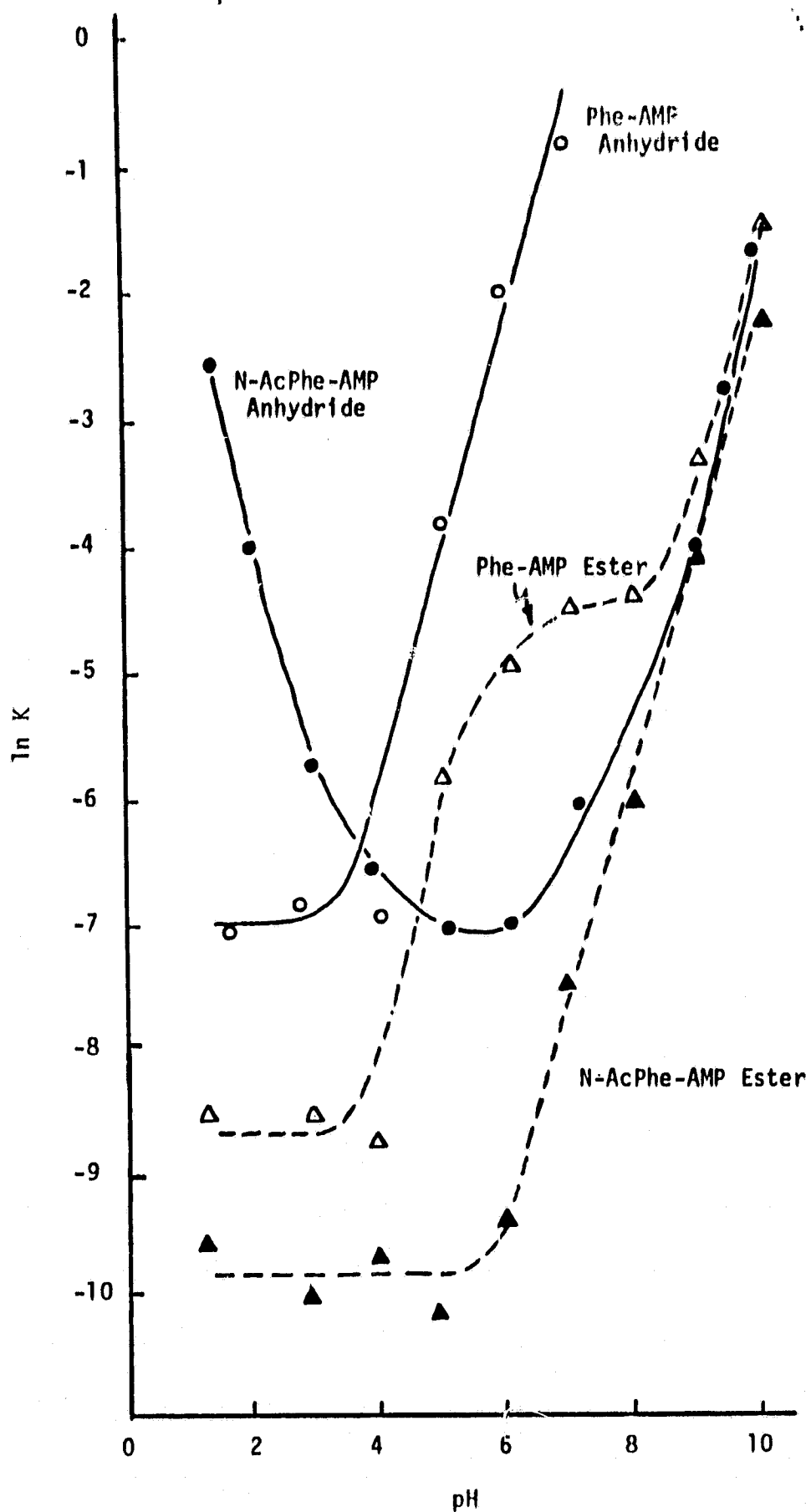


Fig. 6.  
Plot of the natural log of the rate constants for disappearance of the active species at 25°C as a function of pH. Reaction conditions and assay procedures are described in text.

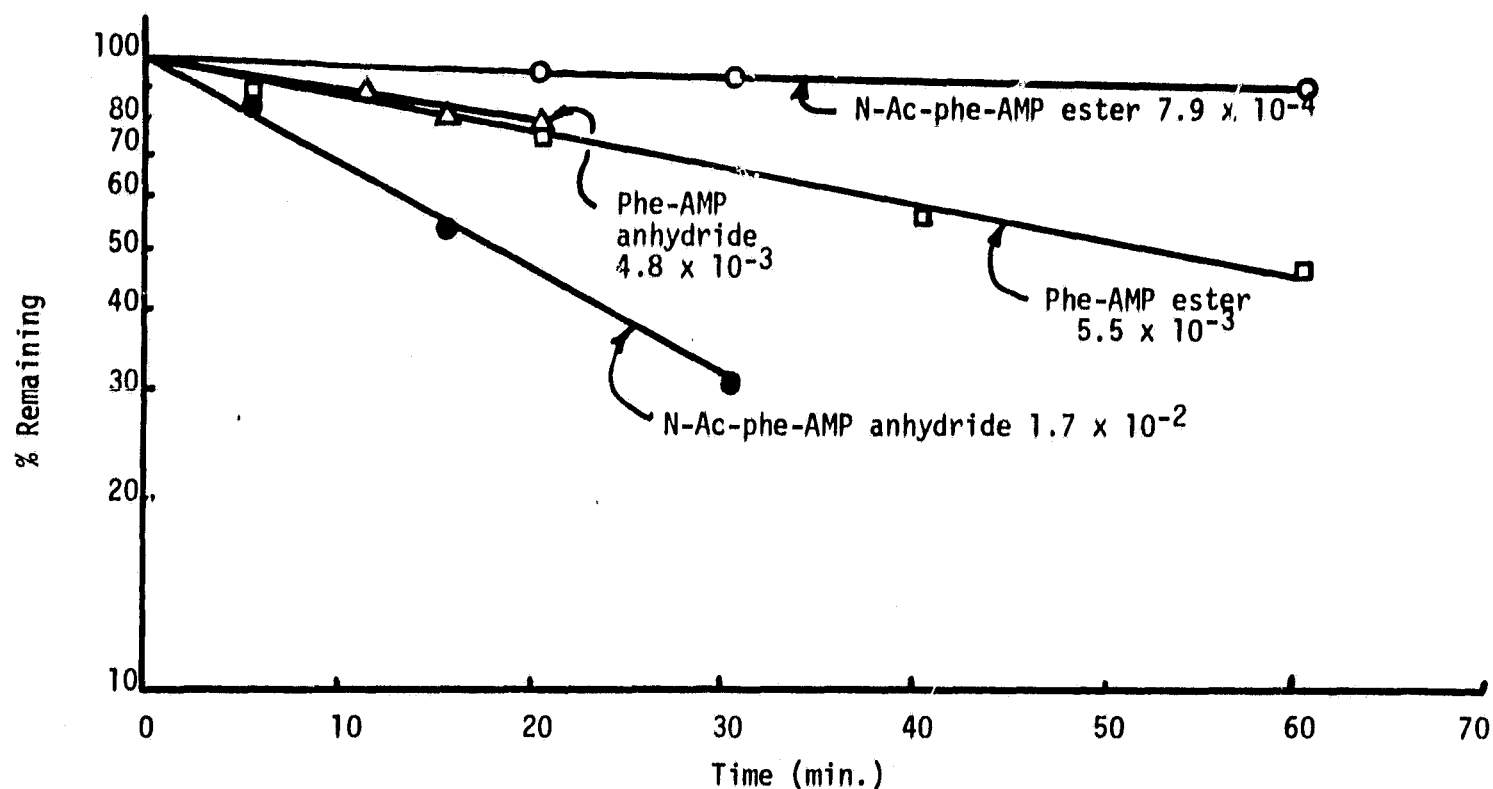


Fig. 7. The % active species remaining when incubated at 25°C, pH 5.5 in 0.25 M hydroxylamine. The pseudo first order rate constants ( $\text{min}^{-1}$ ) are shown along with the identification of the lines. The essential reaction is the formation of hydroxamate from the active species. The rate constants are  $\text{min}^{-1}$ .